IBM “MXT” Memory Compression Technology Debuts in a ServerWorks Northbridge

R. Brett Tremaine

Senior Technical Staff Member

IBM, TJ Watson Research Center

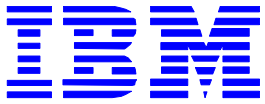
Yorktown Heights, NY

ServerSet III and MXT Technology

Authors

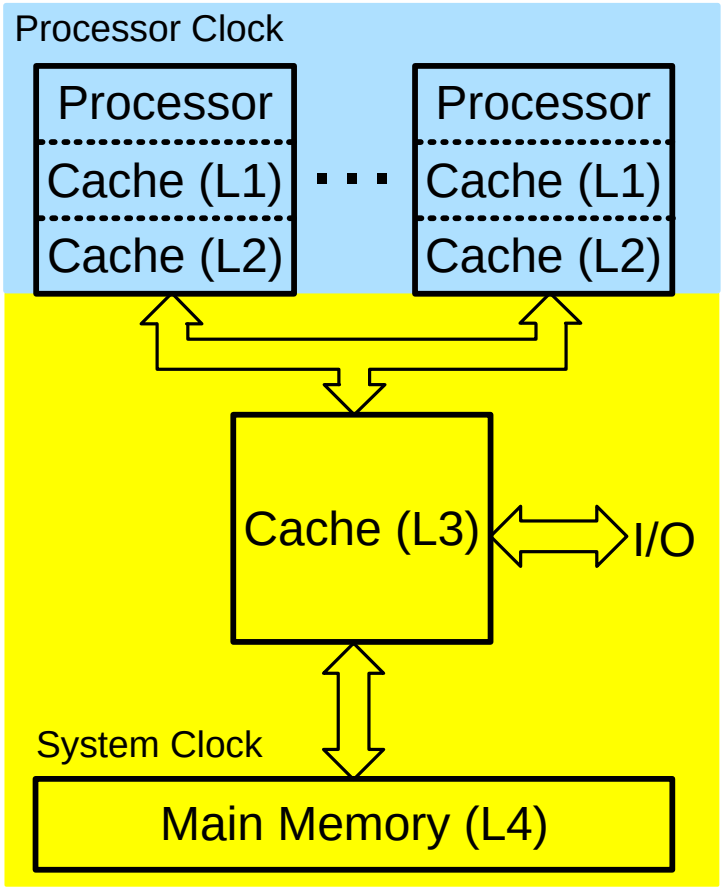
Sujith Arramreddy
ServerWorks
2251 Lawson Lane
Santa Clara, CA 95054

David Har, Kwok-Ken Mak (now at Cisco), R. Brett Tremaine, Michael Wazlowski
IBM, TJ Watson Research Center
Yorktown Heights, NY 10598



ServerSet III and MXT Technology

System Architecture



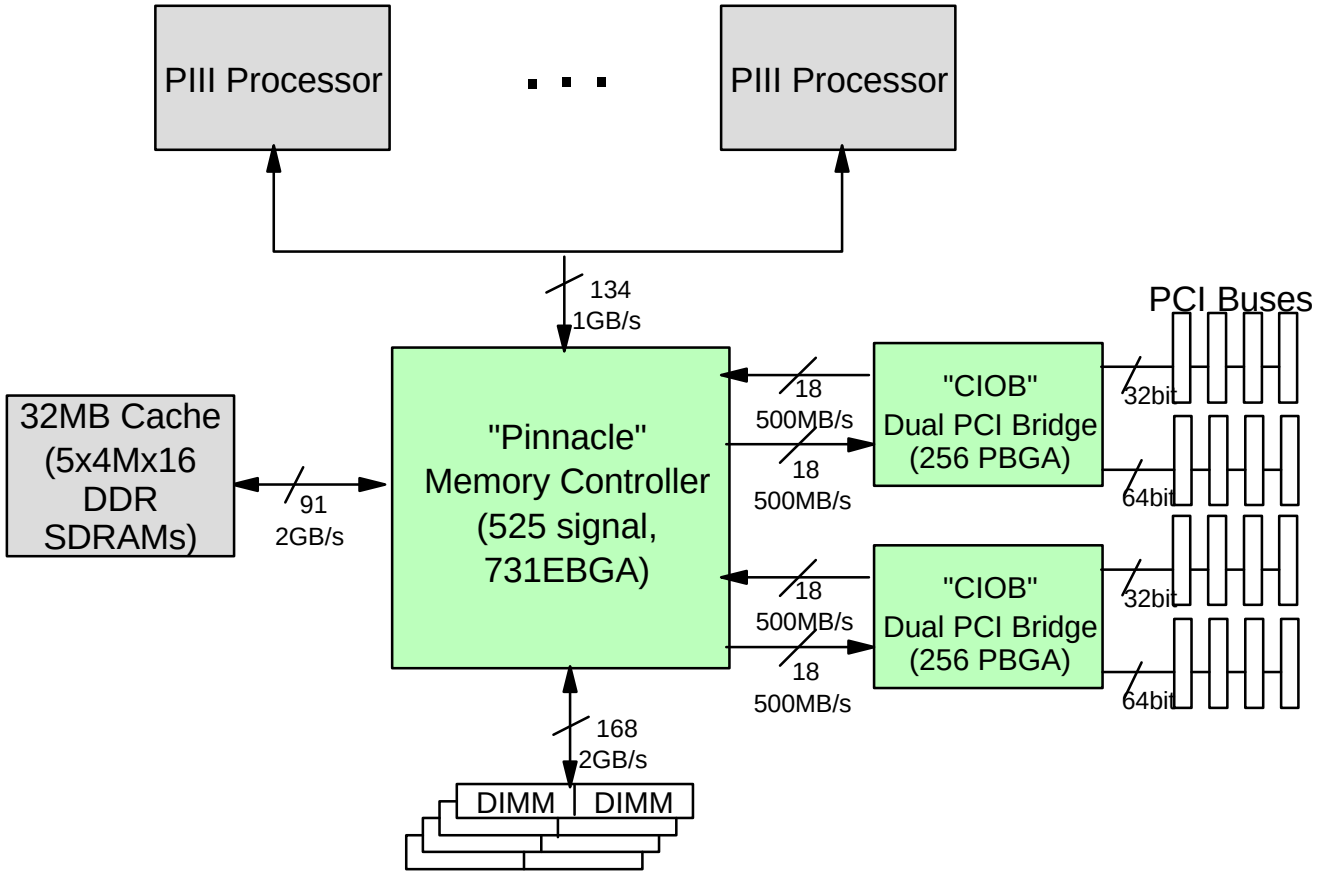
- 2-Level main memory architecture (L3/L4)
  - Fast SDRAM L3 Cache (25% - 40% faster (lower latency) than main memory)
  - L3 Cache Miss Rate < 5% (Typically 2%)
  - High Function L4 Main Memory (Sub-SystemCompression, Online Maintenance, Remote/Very Large Memory)
  - Function Costs Approximately $50 - $60
- Architectural optimization around contemporary technologies
  - 0.25 micron and smaller CMOS
  - Low-Cost high-Density packaging
  - DDR SDRAM
- Cost and performance competitive single chip memory controller for the high volume server and work station market.
  - First of it kind to employ real-time hardware compression to effectively double usable memory
  - Maintains infrequently accessed memory content stored in an efficient compressed format.

IBM

ServerWorks

ServerSet III and MXT Technology

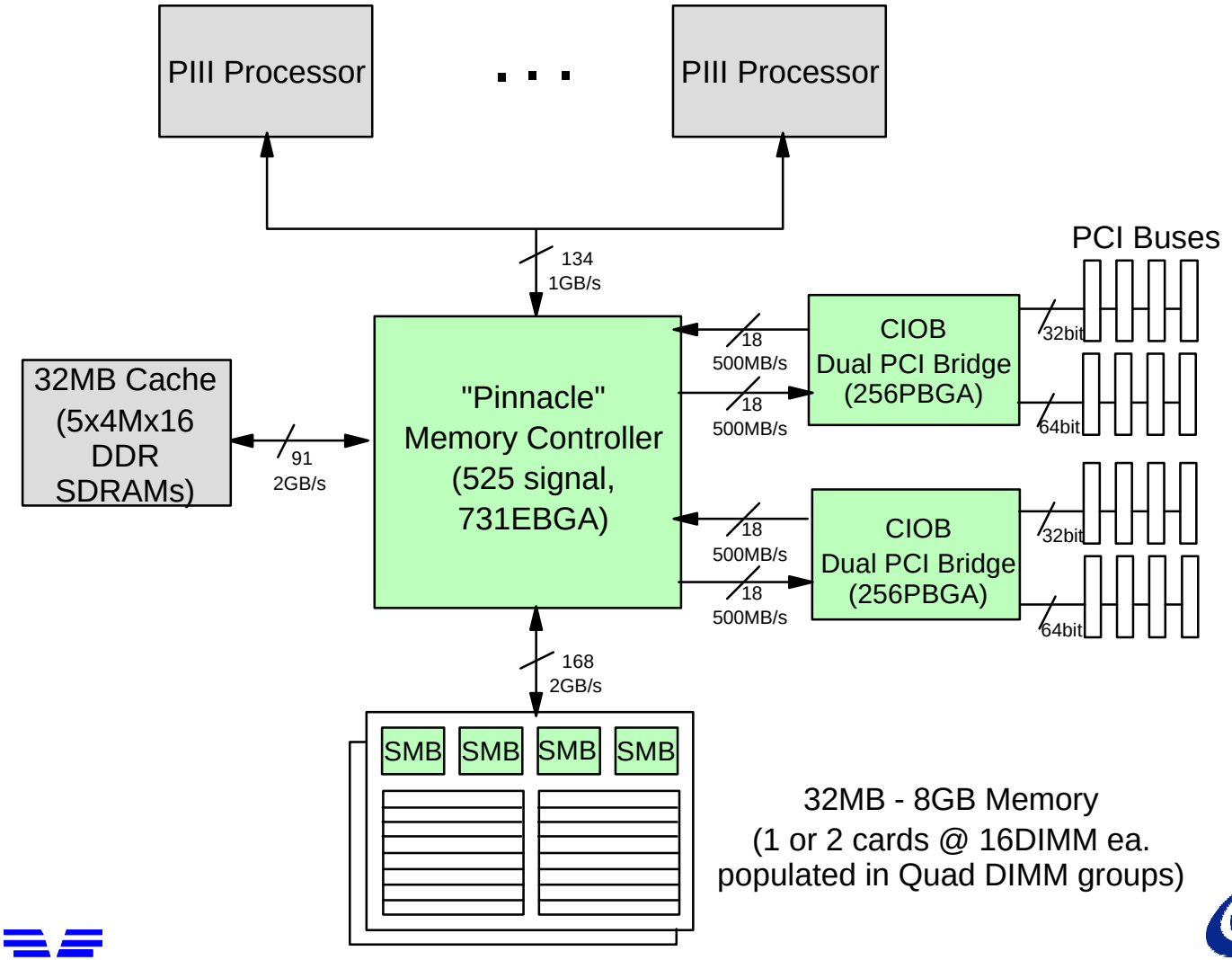
Chip Set Memory Configurations



64MB - 8GB Memory
(4 single/double density SDRAM DIMM's populated in pairs)

ServerSet III and MXT Technology

Chip Set Memory Configurations



32MB - 8GB Memory
(1 or 2 cards @ 16DIMM ea.
populated in Quad DIMM groups)

ServerSet III and MXT Technology

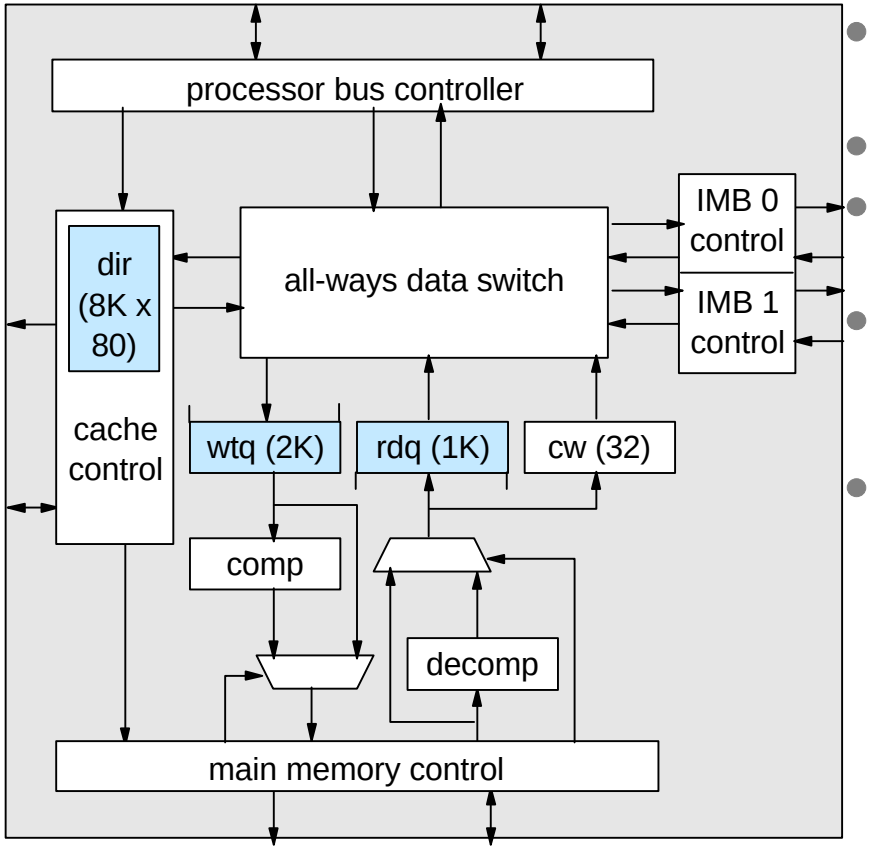
CNB30HE Features

- Single Chip: 525 signal, 731 EBGA package, 2.5VDC power, 100MHZ-133MHz operation
- System Bus Interface
  - **Pentium III Bus (1-4 Processors, full 36-bit address)**
  - **8 Entry Request Queue**
  - **8x32B Memory Write Buffer and 8x4B I/O Write Buffer**
- I/O Bridge Interface (ServerWorks IMB)
  - **Dual Independent Full-Duplex 500MB/s Remote Bridge links**
- Cache Controller
  - **Dual Ported On-Chip Directory (parity protected) Supports 16GB "Real" Addresses**
    - **32MB Cache with 1KB Cache Line Size, 4-Way Set Associative**
    - **Reference State For Snoop Filtering (256B granularity and IO)**
  - **1.6GB/s - 2GB/s access to external 32MB SDRAM DDR data cache (ECC protected)**
  - **Unified data/code/IO memory reference**
  - **LRU Replacement, Write-Back and Write-Allocate Policies**
- Memory Controller
  - **16GB (uncompressed) or 8GB compressed physical memory**
  - **1.6GB/s - 2GB/s Access to 8GB SDRAM Array (ECC Protected for x4 and x8 chip kill)**
  - **Hardware Memory Manager Unit**
  - **Hardware Data Compressor/Decompressor (16B/cy Decompression at 2x Clock and 4B/cy compression) for 1:1 - 64:1, real-time 1KB physical memory block compression**
  - **Programmed 4KB page operations for 1 microsecond speed move, swap, and clear**
- Hardware Performance Monitor
- I2C Bus Interface Access to All Internal Registers

ServerWorks

ServerSet III and MXT Technology

CNB30HE Internal Architecture



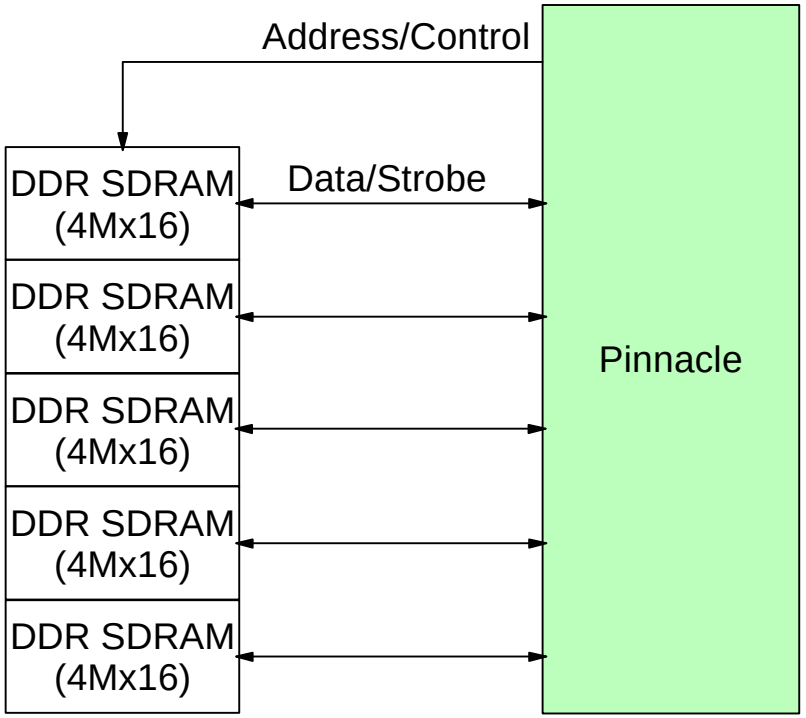
- Fully pipelined 16B internal data flow at system clock frequency
- 2X system clock decompressor.
- 2 outstanding cache miss requests with writeback.
- Early cache miss "data ready" for pre-arbitration of "deferred" processor bus request.
- Performance
  - Write 6 1 1 1
  - Read, Hit Cache (SDRAM row open and hit) 8 1 1 1
  - Read, Hit Cache (SDRAM row open and miss) 14 1 1 1
  - Read, Hit Cache (SDRAM row closed 14 1 1 1
  - Read, Hit Cache (autoprecharge mode) 12 1 1 1
  - Read, Miss Comp (avg.) 69 1 1 1
  - Read, Miss Non-Comp 24 1 1 1
  - Read, Miss Comp. Off 16 1 1 1

IBM

ServerWorks

ServerSet III and MXT Technology

Shared Cache



- 32MB (32K x 1024B lines), 4-way set associative cache, ECC protected
  - 5 x 4M x 16 DDR SDRAM array (burst 4)
  - 8K x 80 dual port SRAM on-chip directory
- 256B coherency granularity with IO "presence" bit.
- Special provsion for two cache lines to exist within one physical line during castout.
- Independent line-fill and writeback controllers track one-another with CW fill order.
- Processor priority access to cache.